We document a decline in the frequency of shopping trips in the United States since 1980 and consider its implications for the measurement of consumption inequality. A decline in shopping frequency as households stock up on storable goods (i.e., inventory behavior) will lead to a rise in expenditure inequality when the latter is measured at high frequency, even when underlying consumption inequality is unchanged. We find that most of the recently documented rise in expenditure inequality in the United States since the 1980s can be accounted for by this phenomenon. Using detailed micro data on spending, which we link to data on club/warehouse store openings, we directly attribute much of the reduced frequency of shopping trips to the rise in club/warehouse stores. (JEL D12, D31, D63, D91, E21)

Income inequality has been rising sharply since the 1980s, raising concern among economists, policymakers, and the general public. However, whether consumption inequality has gone up in similar fashion, which is arguably more relevant for welfare, remains the subject of heated debate in the literature. Understanding what has happened to consumption inequality can also be informative about the forces underlying the rise in income inequality.

For example, Krueger and Perri (2005) argue that improved financial intermediation has allowed households to more easily smooth their consumption over transitory income shocks, thereby compressing consumption inequality. Relatedly, Blundell, Pistaferri, and Preston (2008) argue that much of the rise in income inequality since the mid-1980s came from transitory shocks (as opposed to permanent shocks) that households are able to partially insure themselves against, consistent with the Permanent Income Hypothesis (PIH) and the absence of a commensurate rise in consumption inequality. On the other hand, Battistin (2003), Attanasio, Battistin, and Ichimura (2007); Attanasio, Hurst, and Pistaferri (2015); Aguiar and Bils...
(2015); and others argue that the flat profile of consumption inequality is nothing more than a measurement artifact, and that consumption inequality has risen in line with income inequality since 1980. This line of research highlights the value of measuring spending precisely over relatively short periods and thus motivates the use of “big data” with large cross sections and short time series.

We build on this literature by emphasizing the distinction between spending (expenditure) inequality and consumption inequality. While households can enjoy a smooth consumption flow from most goods, their purchases may occur only infrequently. Because household surveys typically track expenditures for a short period of time to minimize recall error and reporting burden (e.g., a two-week period in the Diary Survey of the Survey of Consumer Expenditures (CEX)), measures of spending inequality can fail to correctly measure the underlying consumption inequality due to the timing of purchases. This can matter not only in the cross section (if one household happened to buy paper towels in a period and another did not, spending inequality over that period would be high even if both households have the same flow consumption of paper towels) but also for measuring trends over time. To see the latter, suppose that consumers start stocking up on food once a month rather than once a week. Even if they maintain the same consumption flow, the cross-sectional inequality of spending measured at a less than monthly horizon will rise despite the fact that underlying consumption inequality would have stayed the same. In this paper, we document such a decline in the frequency of shopping, quantify its potential implications for historical changes in consumption inequality, and study its potential sources.

Our starting point is the well-documented difference in the trend of expenditure inequality across the Diary and Interview surveys of the CEX over the 1980s and 1990s (Bureau of Labor Statistics (BLS), 1980–2015). While the Interview Survey points toward little change in expenditure inequality over time (as documented in Krueger and Perri 2005), the Diary Survey instead suggests that expenditure inequality has risen more closely in line with income inequality (see, e.g., Battistin 2003). Although there are many potential sources for this difference, one is the differing frequency over which expenditures are measured: biweekly in the Diary Survey and monthly (or quarterly for some categories) in the Interview Survey. While a rise in weeks with “zero” spending in the Diary Survey has been documented before (Bee, Meyer, and Sullivan 2015), whether this rise is the result of increasing nonresponse or a decline in the frequency of shopping (“true” zeros) has not yet been investigated in detail. Consistent with the frequency of expenditure measurement being a force behind the different inequality trends in the two surveys, we document that the frequency of shopping has indeed systematically declined from the 1980s to the present day. Using data from the CEX Diary Survey, we find that the fraction of days in which households engage in any shopping for nondurable goods has been falling over time, so that households concentrate their shopping into fewer days of the week. Using even more detailed information on household expenditures from the Nielsen HomeScan data, we again document a decline in the number of days

1 Attanasio et al. (2007) provide a discussion of other potential sources for differing trends in expenditure inequality across the two CEX surveys.
in which households do their shopping since 2004, consistent with what we see in
the Diary Survey over this period (Nielsen Company 2004–2014). Hence, part
of the greater increase in inequality as measured by the Diary Survey may indeed be
coming from a changing frequency of shopping by households.

We provide several additional pieces of evidence, based on micro data, consistent
with this conjecture. First, while average real expenditures on goods in the Nielsen
sample has been approximately constant between 2004 and 2014, this stability
masks underlying changes along the intensive and extensive margins of shopping
behavior. The number of shopping trips (extensive margin) has been steadily falling
over the entire sample, whereas the average expenditures per trip (intensive margin)
have been rising. Hence, we see households making fewer, but larger, shopping trips
on average. Second, using information on the volumes and sizes of individual goods
purchased in the Nielsen sample, we find that households have been purchasing
larger quantities or volumes of goods over time, consistent with increased stocking
up. Third, using the American Time Use Survey, we compute average shopping
times for individuals (Bureau of Labor Statistics (BLS), 2003–2015). We find a
strong decline in the average amount of time spent shopping by US households,
driven entirely by the extensive margin. Households do fewer trips per day and are
less likely to go to any store on any given day. In contrast, the average duration of a
shopping trip (the intensive margin) has held steady over this time period. These are
precisely the expected patterns as households buy larger quantities of goods while at
the store and therefore need to go to the store less often. Thus, the ability to stock up
appears to be a critical component of these differences in trends.

To quantify the contribution of changing frequency of shopping to the differential
trends in expenditure inequality across frequencies of aggregation, we pursue
two approaches. Our first approach is to examine trends in inequality by the fre-
quency in which expenditures are aggregated. Ideally, one would simply vary the
duration of periods over which each household’s expenditures are aggregated then
construct cross-sectional measures of dispersion for each frequency. Unfortunately,
the data in each of the CEX surveys are inadequate for this since households in the
Diary Survey only report their expenditures for two weeks, while households in the
Interview Survey report their expenditures over one-to-three month periods but do
not provide higher frequency variation within those periods. The Diary Survey and
Interview Surveys also have different collection methodologies and the same indi-
viduals are not linked across the two surveys.

To examine aggregation biases in the same dataset, we rely on the Nielsen
Homescan data, which are available at the level of a shopping trip. The Nielsen
data track spending by households daily for extended periods, thereby allowing us
to assess the extent to which trends in expenditure inequality are sensitive to the
frequency over which expenditures are aggregated, e.g., biweekly versus monthly.
Ideally, we would like to get a sense of the size of the aggregation bias going
back to the 1980s and 1990s when the DS and IS diverge most prominently, but

---

2 We first perform a battery of checks to ensure that the Nielsen data are comparable to CEX Diary Survey. We
find that mean expenditures and implied inequality levels are quite close across the two datasets, once one focuses
on goods that are common across the two.
unfortunately Nielsen data are only available from 2004. Still, this approach yields a clear effect of time aggregation on the trends in spending inequality. Short time horizons for measuring consumption imply positive trends in inequality, whereas lower frequencies of aggregation imply much flatter profiles of inequality. Hence, time aggregation can effectively account for all of the difference in the trends of consumption inequality identified by these two surveys.

Our second and complementary approach to quantifying these effects exploits the changing dispersion in individual households’ expenditures over time (i.e., dispersion of a household’s expenditures over the course of a year when expenditures are measured weekly, biweekly, etc.). We show, using a simple model, that the latter measure provides a convenient metric to assess the contribution of shopping frequencies to the cross-sectional dispersion of expenditures. While the Nielsen data are available only for 11 years, these data are most informative for linking time-series and cross-sectional dispersion of spending. Based on our analysis of the Nielsen data, we can attribute all of the differential increase in cross-sectional expenditure inequality across the two surveys to a changing frequency of shopping. Jointly, these results therefore suggest that the Interview Survey generates more reliable measurement of actual changes in consumption inequality than the Diary Survey, which supports the Kruger and Perri (2005) argument, based on IS data, and so we conclude that consumption inequality has generally not kept pace with income inequality since the 1980s.

There are many mechanisms that could explain why American households are increasingly purchasing larger quantities when shopping and therefore shopping less frequently. One such mechanism is the rise of club/warehouse stores (Costco, Sam’s Club, BJ’s, etc.) which, by design, sell larger quantities of goods to households at lower unit prices. As these stores have expanded throughout the country since the 1980s, it has become easier for households to stock up in ways that were not feasible in the past, consistent with the decreased frequency of shopping that we observe. Furthermore, there is considerable geographic variation in the ease with which households can access one of these retailers, enabling us to quantify the contribution of this mechanism.

To assess whether club/warehouse stores can explain some of the rising concentration in household shopping trips, we again use the unique richness of the Nielsen data to characterize the link between how much variation there is in an individual’s spending over time and their reliance on club stores in their expenditures. Specifically, we first measure the variation in a household’s expenditures over

---

3 We perform a similar test in the CEX Diary Survey by comparing trends in inequality of expenditures summed at the weekly versus biweekly frequency and in the CEX Interview Survey by comparing trends in inequality of expenditures summed at the quarterly versus annual frequency. In each case, we find the same qualitative result that higher frequencies of aggregation lead to steeper trends in expenditure inequality.

4 The use of IS versus DS data is not the only point of contention in this literature, however. For example, papers using the PSID tend to find a rise in consumption inequality since the mid-1970s, although less than for income inequality (Attanasio and Pistaferri 2014). Prior to 1994, the PSID asked about annual food expenditures; since 1994 households were allowed to report food spending for higher frequencies (e.g., 1 or 2 weeks), and spending is then annualized. This recent spending in the PSID is therefore susceptible to the same changes in shopping patterns. To the extent that households use heuristics before 1994 to answer questions (e.g., annualizing last week’s spending), changes in shopping patterns could also play a role.
a year using different time frequencies for subperiods. Households who do more infrequent shopping trips have relatively higher dispersion in their expenditures at higher frequencies than at lower frequencies. This greater time-series dispersion in expenditures for one household when it does large purchases infrequently is therefore analogous to how cross-sectional dispersion in consumption is higher when more households engage in infrequent shopping.

To assess how much households use club stores, we measure the fraction of a household’s expenditures that were spent at a club store over the course of that year. The link between shopping at club stores and stocking up can then be assessed by regressing an individual’s expenditure dispersion on that individual’s share of expenditures going to club stores, using dispersion measures at different time frequencies. The results suggest that shopping at club stores is indeed correlated with significantly more stocking up. There is a strong positive correlation between the coefficient of variation at the weekly frequency and a household’s share of expenditures at club stores, but this correlation declines rapidly as we increase the amount of time over which expenditures are aggregated, as expected since it becomes progressively more difficult to stock up for longer periods. Shopping at club stores also explains a diminishing fraction of the variance in households’ coefficients of variation at longer durations for time aggregation of expenditures. An instrumental variable strategy, based on the differing distance of households from club/warehouse stores, supports causality running from access to club/warehouse stores to increased stocking up in expenditures. We use these estimates to impute the contribution of club stores on the rise in measured inequality in the Diary Survey. We find that the increased prevalence of club/warehouse stores since the early 1980s can account for approximately 30 percent of the rise in measured inequality in expenditures.

This paper relates to a growing literature on measuring and understanding the sources of economic inequality. Unlike much recent work on the rising share of income and wealth of the top 1 percent (e.g., Piketty and Saez 2003, Saez and Zucman 2016), we focus on inequality outside of the top 1 percent since our data sources are not informative about top income earners. Instead, our results build on the literature relating consumption and income inequality amongst households in the bottom 99 percent in the United States (e.g., Krueger and Perri 2005, Aguiar and Bils 2015, Attanasio and Pistaferri 2016, Meyer and Sullivan 2017) or abroad (e.g., Gorodnichenko, Peter, and Stolyarov 2010). Relative to these papers, our contribution is to document how the changing frequency of shopping contributes to the measured trends in spending inequality. This paper is also closely related to a growing literature on shopping behavior of households and its implications for macroeconomics. For example, Nevo and Wong (2016) focus on the substitution between expenditures and home production during the Great Recession. Wong (2015) studies infrequent purchases of durable goods, Aguiar and Hurst (2013) focus on life-cycle consumption patterns of households, and Coibion, Gorodnichenko, and Hong (2015) study the implications of store-switching for macroeconomic dynamics. We extend this line of work to study how shopping behavior can affect the measurement of expenditure inequality. Another closely related line of work focuses on household search for goods. Menzio and Trachter (2015), for example, focus on the implications of consumer search for equilibrium price dispersion, while Michaillat and Saez (2015) study the implications
of product market search for macroeconomic dynamics. Finally, our use of time use data to study household decisions echoes Aguiar, Hurst, and Karabarbounis (2013) and Lee, Kawaguchi, and Hamermesh (2012). But whereas they use time use surveys to study the substitution between labor, leisure, and home production, we exploit information on time spent shopping and traveling to stores to characterize the changing nature of household shopping behavior in the United States.

The paper is organized as follows. In Section I, we discuss the different CEX surveys and their implications for trends in expenditure inequality, as well as providing preliminary evidence on how the frequency of shopping has declined over time. Section II introduces the Nielsen data and provides additional evidence on the changing characteristics of household shopping. Section III provides two ways of quantifying the contribution of these changing shopping patterns to expenditure inequality trends. Section IV assesses how much of the changes in household shopping behavior can be attributed to the growing prevalence of club stores. Section V concludes.

I. Expenditure Inequality and the Changing Frequency of Purchases

To measure consumption inequality, previous work such as Krueger and Perri (2005) and Meyer and Sullivan (2017) has focused primarily on the Consumer Expenditure Survey (CEX). Since the CEX is a well-known and well-documented data source, we provide only a brief overview of these data. We focus in particular on the differences between the two main components of the CEX, both of which have been used to measure consumption inequality: the Interview Survey (IS) and the Diary Survey (DS). We also highlight changes in the survey methodology over time that could impact the dispersion of measured spending.

In the IS, about 1,500–2,000 households are asked each month to recall the dollar value of spending over the previous month or quarter (depending on the category). Households are interviewed once per three months for five consecutive three-month periods, although the BLS only makes data available for interviews two through five. While early Interview surveys exist in 1960–1961 and 1972–1973, the modern Interview Survey begins in 1980 and is not directly comparable to prior waves of the IS. In the 1982–1983 waves for the IS and 1982–1983 waves for the DS, only urban households were sampled due to budget cuts. A main advantage of the Interview Survey is its broad coverage of goods purchased by households (approximately 95 percent of typical household’s consumption expenditures) since it is used to create expenditure weights for the Consumer Price Index.

A separate sample of households participate in the Diary Survey. Households are asked to record their spending each day for two weeks in a diary, which is later transcribed by US Census Bureau officials. Records of daily spending become available to researchers starting in 1982, for the categories of food-at-home as well as food away from home. In 1986, the Diary Survey was expanded to cover a more comprehensive set of spending categories. In 2004, the Census Bureau adopted a variety of changes to Diary data collection procedures that resulted in potentially more accurate recording of purchases, including computer assisted technology for US Census Bureau enumerators. In the figures made using the DS, we include a vertical line to indicate these structural breaks.
An extensive literature exists discussing the pros and cons of the two surveys. For example, Krueger et al. (2010); Aguiar and Bils (2015); Attanasio, Battistin, and Ichimura (2007); and Battistin and Padula (2016) find that the Interview Survey in the CEX underreports spending relative to aggregate data and that this underreporting has become more severe over time. On the other hand, Bee et al. (2012) compare reported consumption spending data in the CEX to comparable data from the national income accounts data and find that the CEX data conform closely to aggregate data for large consumption categories. Battistin (2010) and Attanasio, Battistin, and Ichimura (2007) argue that, given data in the DS, the IS underestimates the rise in expenditure inequality since the 1980s. In contrast to the view promoted by Krueger and Perri (2005) and Meyer and Sullivan (2017) that expenditure inequality (measured using the CEX IS) has not risen nearly as much as income inequality, this alternative view implies that expenditure inequality has in fact grown more rapidly than implied by the CEX IS.

To illustrate how pronounced the differences between the Interview and Diary surveys are for resulting trends in expenditure inequality, we construct a coefficient of variation for each survey. Specifically, for each survey, we measure each household’s expenditures on nondurable goods and services. In the Diary Survey, expenditures are daily while in the Interview Survey they are over a monthly or quarterly horizon. We use BLS’s monthly Personal Consumption Expenditures (PCE) price index to deflate household spending (for quarterly spending, we follow the BLS and allocate spending equally across months). Because the CEX IS and DS surveys use a somewhat different classification of goods and services (universal category codes, UCCs), we update the concordance created in Bee et al. (2012) and increase its scope to be comprehensive of all UCC codes (see online Appendix E). We then calculate the coefficient of variation in expenditures across all households (the ratio of the cross-sectional standard deviation to the cross-sectional mean of expenditures) for each year.

We use the coefficient of variation (CV) as our baseline measure of inequality because it allows us to include households reporting zero spending in a given period. The latter is an important constraint because, as we document below, it is common for households to report zero weekly (or biweekly) spending for the studied categories of goods. Other popular measures of inequality (e.g., 90/10 ratio, standard deviation of logs) are ill-suited to the presence of many zeros in the data and therefore would not provide a comparable measure of inequality for the highest-frequency (e.g., weekly) data on spending and subcategories of goods with infrequent purchases. In addition, our model allows us to derive a simple relationship between the time-series and cross-sectional coefficients of variation—whereas no such simple formulas exist for other measures of inequality such as the Gini coefficient.

5 Some of the differences between CEX and aggregate data could be due to the fact that these two sources use increasingly different definitions and coverage. For example, aggregate data include purchases by non-profitable organizations and business which have grown as a share of the national accounts over time.

6 Our coverage of nondurable goods and services follows Coibion et al. (2017). Clothing and most services are not consistently measured in the Diary survey until 1986. To minimize any adverse effects of outliers on measures of inequality, we winsorize the right tail of household spending for a given frequency in each year at 1 percent.
The resulting time series are plotted in Figure 1 (bootstrapped standard errors for the CV series are reported in online Appendix Figure 2). Using the Interview Survey, we replicate the baseline result of Krueger and Perri (2005), finding little increase in expenditure inequality between 1980 and 2015. In contrast, the Diary Survey reveals a pronounced increase in expenditure inequality from 1980 to the early 2000s. The ratio of the two inequality measures provides a simple way of examining differences in trends across the two: this ratio is increasing systematically over time, going from 1.05 in 1980 to 1.35 in 2015. Its persistent increase even since the early 2000s reflects the fact that spending inequality as measured by the Interview Survey is declining over this time period but approximately constant in the Diary Survey. This difference in trends (rather than the trends themselves) is the focus of our analysis.

The diverging trends in inequality across the two survey measures are not driven by composition effects, either in terms of composition of goods or characteristics of households. For the former, we can compare spending inequality in the two surveys for matched and consistently (over time) collected categories of goods, thereby controlling for potential changing compositions of purchases over time. We find that the same trend in the ratio of inequality across the two surveys holds (online Appendix Figure A1). Similarly, we can control for potentially changing household characteristics by looking at residual inequality in each survey. We do so by regressing household expenditures on a large set of observable characteristics of households.
(age, income, etc.) in each survey, then construct equivalent inequality measures from the residuals of household expenditures:

\[(1) \quad Y_{ht} = X_{ht} \gamma + \epsilon_{ht},\]

where \(h\) and \(t\) index households (respondents) and years, \(Y\) is a variable of interest, \(X\) is a vector of controls that includes a polynomial in the age of household head, gender dummy for household head, a set of race dummies for household head, a set of dummies for educational attainment of household head, number of children, dummy for employment of household head, and a set of region dummies. The coefficient of variation adjusted for changes in demographics is calculated as \(\sqrt{\text{var}(\epsilon_{ht}) / \text{mean}(\hat{Y}_{ht})}\). The results yield a similar pattern of a systematically rising ratio of consumption inequality in the DS relative to the IS from 1.1 in 1980 to 1.5 in 2015 (see Figure 1). Results are also similar if we use different metrics for measuring inequality (Gini coefficients and 90/10 ratios are plotted in online Appendix Figure A3) or within subgroups of the population. For example, in online Appendix Figure A4, we document that the same patterns of rising inequality in DS survey relative to IS survey hold within both younger and older households, for the employed and the nonemployed, for different races, for households of different sizes, as well as for households of high and low income.7

It is also worth noting the large difference in level between the two series. Although the IS likely has larger measurement error than the DS, this difference in levels is to be expected since the DS measures expenditures at the biweekly frequency, whereas the IS measures expenditures over a monthly/quarterly horizon. Since some goods are purchased infrequently, the Diary Survey will record zero expenditures for some households and large expenditures for others depending on the timing of their purchases. In contrast, the Interview Survey will more consistently measure positive expenditures due to the longer horizon. By the same logic, inequality among weekly household expenditures in the Diary Survey is approximately 20 percent higher on average than for expenditures at the biweekly frequency in the same survey (see online Appendix Figure A2).

Importantly, the fact that expenditures are measured over different horizons can be a source of differences in trends of measured “consumption” inequality if the frequency of household purchases is changing over time. For example, if households change their frequency of purchasing toilet paper from a biweekly to a monthly frequency while keeping their flow consumption of toilet paper unchanged, this would induce a rise in the ratio of spending inequality when expenditures are measured at the biweekly frequency relative to when expenditures are measured at the monthly frequency. In this case, consumption inequality would not have changed (everyone is still using the same amount of toilet paper per unit of time), but inequality in spending at high frequencies would rise, thus underscoring the difference between spending inequality and consumption inequality.

We have also calculated the CV for income in both of the surveys, and find that income inequality has risen hand-in-hand in both, so these differences cannot be explained by differences in the cross-section of income across the two surveys. See online Appendix Figure A6.
There is evidence consistent with this hypothesis. Since the CEX Diary Survey provides daily expenditures, we can measure the average number of days in which households engage in positive expenditures (out of 14 possible days in the diary) for each survey year. The result is plotted in Figure 2. Panel A refers to all nondurables including services, while panel B focuses on a more limited set of nondurable goods: food at home, alcohol, tobacco, and other small nondurable goods. In 1980, households purchased a positive amount of nondurable goods 9 out of the 14 days of each biweekly period, but this number had fallen to 8 days by 2004. There is a structural break in the series of 2005 (because of the changes in how CEX DS data are collected), but the average number of days falls another 0.5 by 2015. Panel B shows an even steeper drop in shopping for the narrower set of nondurables. Almost identical trends are obtained if we use positive values (e.g., $5, $10, etc. including with inflation adjustment) as the threshold for daily expenditures instead of zero.

Like the changing ratio of expenditure inequality in the two surveys, the declining frequency of shopping, at least as measured in days with positive spending, holds for a wide range of products and is not driven by household characteristics, such as a growing share of working spouses. To see the latter, we also plot in Figure 2 residual measures of the number of days with positive expenditures for each household after controlling for the household observable characteristics, as before, and measure the average across households (normalizing it to have the same value as the raw measure in 1980 and again in 2005). The trends are almost identical, so the declining frequency of days with positive shopping experiences is not coming from changes in household characteristics.

Unfortunately, the CEX data present many limitations which do not allow us to characterize these effects in a more detailed way. For example, without more detailed information on households’ shopping activities, we cannot quantify whether households are doing fewer shopping trips or are combining the same number of trips into fewer days. Without information on quantities and sizes of purchased goods, we cannot assess, e.g., whether households are buying larger quantities on their less frequent trips. Without information on time use, we cannot determine whether households are changing the amount of time they devote to shopping. And because neither the DS nor IS has long panels of high-frequency data on expenditures, we cannot quantify the extent to which changing frequencies for computing expenditure inequality contributes to the differential trends in spending inequality. Using additional datasets, we explore these questions in the next section.

8 Since the Interview Survey does not provide high-frequency expenditure data, we cannot construct equivalent measures in that data.
9 Attanasio and Davis (1996) propose an alternative approach based on synthetic cohorts. In short, one can create many bins of households similar in some dimension, then compute average spending for each bin and finally use average spending to calculate inequality across bins, which should approximate inequality well when the number of bins is large. If there are sufficiently many households in each bin, then whether a given households happens to buy a good or not (because of infrequent purchases) is less important. Intuitively, to compute consumption flow for a given household bin, one may use either a long history of spending for a household in the bin or a large cross-section of households in the bin. The main constraint in this approach is the number of households in a bin. For example, the IS has about 6,000 participating households per year and so the size of 1-percent bin could be too small (approximately 60 households) to average out idiosyncratic variation in the timing of shopping trips.
Panel A. All nondurables

Panel B. Food at home, alcohol/tobacco, and small nondurables

Figure 2. Frequency of Shopping, CEX Diary Survey

Notes: The figure plots the average number of days in which households report any positive spending in CEX (measured for two-week periods) over time. The red line with circles is residual measures after controlling for observable household characteristics, as described in Section I. The vertical dashed line shows the time when the CEX Diary Survey had a change in how it collects data. See Section I for more details. In panel A, the sample is all nondurables (see online Appendix B). In panel B, the set of goods is restricted to be common to both the CEX surveys and the Nielsen data. The included categories are food-at-home, alcohol/tobacco, and small nondurables, which matches the coverage of goods in our Nielsen sample.
II. Changing Patterns of Household Shopping Behavior

Evidence from the CEX Diary Survey suggests that households have been reducing the frequency of their shopping over time, a feature potentially explaining the differing trends in spending inequality observed in the Interview and Diary Surveys. In this section, we provide more detailed microeconomic evidence on the changing patterns of household shopping behavior. To do so, we begin by introducing an additional but more detailed dataset on household shopping, the Nielsen Home Scanner data, and show that these data display similar shopping characteristics as the CEX. For example, in addition to presenting comparable measures of expenditures on different categories of goods, we confirm the finding from the CEX that households are doing their shopping on fewer days using the Nielsen data. Second, we show that while average real expenditures by households were relatively constant between 2004 and 2014, households have been spending relatively more on each shopping trip but doing fewer shopping trips per year, consistent with increased stocking up. Third, using data on quantities, we show that households are purchasing ever larger sizes or numbers of units of individual products on each trip. Fourth, we turn to time use surveys to show that the amount of time households have been spending on shopping has been decreasing. All four facts are consistent with households increasingly “stocking up” in their shopping.

A. Characteristics of the Nielsen Data

While CEX allows us to construct time series going back to the early 1980s and have a good coverage of goods and services purchased by households, the data in each of the CEX surveys present some limitations. For example, households in the CEX Diary Survey only report their expenditures for two weeks, so we may be missing important expenditures that are not made over that two-week measurement period. Households in the CEX Interview Survey report their expenditures over one month (or three months depending on the category) but do not provide higher frequency variation within those periods. Because the Diary and Interview surveys are not connected in any way, we cannot establish how time aggregation affects trends in measured spending inequality. Moreover, measurement error due to recall bias in the IS imposes additional challenges when comparing consumption inequality at quarterly frequency in the IS versus higher frequency in the DS.

To address these challenges, we turn to Nielsen Home Scanner (Nielsen) data, available through the Kilts Center at the Booth School of Business at the University of Chicago, which provides a source of rich, high-frequency household spending data. Nielsen data are currently available from 2004 to 2014. From 2004–2006, the sample included approximately 40,000 households, increasing to 60,000 households beginning in 2007. Over the period 2004–2014, the mean and median tenure in the sample were approximately four and three years, respectively.¹⁰

¹⁰Households participating for a long time in the Nielsen panel may exhibit fatigue in reporting their purchases and shopping trips. We found that controlling directly for tenure in the panel does not materially affect the moments.
The Nielsen sample is comprised of a combination of households recruited by Nielsen, as well as unsolicited volunteers. In exchange for their participation, households receive points that can be redeemed for prizes as well as entry into lotteries that award more points or cash. Households are provided a scanning device by Nielsen to scan the barcodes of their purchases, and they are encouraged to scan newly purchased items as soon as they return home. Nielsen employs their own sample filter, requiring that households must report a minimum dollar amount per month, which varies depending on household size, to be in the final sample. To ensure that our results are not driven by households with incomplete records, we include only households with at least one shopping trip where they scan items in each month of a given year.

After scanning a product using the device, households directly report the quantity of the barcode (or universal product code, UPC, a precise definition of a good) that they purchased. For a group of participating stores, prices are automatically reported to Nielsen; otherwise the household is also asked to manually enter the product price. Nielsen later merges in information about the product that is tied to the barcode, including a measure of volume or count if applicable. If a product does not have a barcode, a purchase of this product is generally not reported as the main Nielsen data focus primarily on nondurables with a barcode.

Household demographics, including zip code and employment status, are updated once per year as part of a household survey. Nielsen uses the demographic information to construct household weights that weigh the sample to be nationally representative. The household spending data are technically available on a daily basis. However, in some cases, the purchase date in Nielsen could reflect the date the data were transmitted by the scanning device to Nielsen, rather than the true purchase date by the household.

The Nielsen data include over 325 million barcodes that Nielsen estimates to cover approximately 30 percent of household spending. These barcodes are categorized by Nielsen into lower levels of aggregation. Nielsen’s “Product Groups,” of which there are 125, are closest to universal classification codes (UCCs) in the CEX Diary Survey. For our analysis, we construct a correspondence table between CEX Diary UCCs and Nielsen Product Groups (see online Appendix B).

To ensure that our results are not driven by the specifics of how Nielsen data are collected, we compare basic moments for categories of consumer spending in CEX Diary Survey and in Nielsen. All statistics are for the biweekly frequency. We compute moments for categories of goods present in both sources so that the coverage of

---

11 Nielsen data are best at capturing goods with universal product codes (UPCs). One concern is that non-barcoded goods could be trending differently across lower and higher income households. Since 2007, a subset of panelists report spending on non-barcoded goods, known as “magnet” products. When we examine spending on these magnet products across income groups, we indeed find large differences in levels of spending on magnet products across income groups, but we do not find evidence that spending on magnet goods is trending differently, at least since 2007.

12 Nielsen made changes in 2009 that resulted in more purchases being assigned a transmission date rather than the true purchase date. We therefore must be cautious comparing higher frequency (especially daily) behavior across these regimes.
goods is comparable across datasets (e.g., Nielsen data have virtually no coverage of services). The set of comparable nondurable goods generally includes food, alcohol, and small nondurables. To differentiate the frequency of shopping trips and the size of purchases, we show the share of households reporting zero spending over two weeks for a given category of goods (“zero share”) and moments (mean, standard deviation, interquartile range) for the size of purchases conditional on a purchase in the category. Results for selected categories of goods for year 2014 are reported in Table 1.

Consistently across data sources, we observe that purchases for many categories of goods are not made frequently. On average, there is an approximately 70 percent chance that there is no purchase in a typical category of goods over two weeks. Furthermore, for the comparable categories, the probability of no purchases for any of the categories during the period is 7 percent in the CEX data and 10 percent in the Nielsen data. The correlation of zero shares across the surveys is 0.74, thus, indicating high consistency across data sources.

Table 1—Biweekly Spending in the CEX Diary Survey and Nielsen Data

<table>
<thead>
<tr>
<th>Spending category</th>
<th>CEX Diary</th>
<th></th>
<th></th>
<th></th>
<th>Nielsen</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard deviation</td>
<td>IQR</td>
<td>Zero share</td>
<td>Mean</td>
<td>Standard deviation</td>
<td>IQR</td>
<td>Zero share</td>
</tr>
<tr>
<td>Total spending</td>
<td>1,475.25</td>
<td>1,576.54</td>
<td>1,519.93</td>
<td>0.00</td>
<td>161.16</td>
<td>125.04</td>
<td>151.00</td>
<td>0.10</td>
</tr>
<tr>
<td>Baby food</td>
<td>19.41</td>
<td>27.90</td>
<td>11.40</td>
<td>0.95</td>
<td>19.95</td>
<td>27.94</td>
<td>16.28</td>
<td>0.98</td>
</tr>
<tr>
<td>Pet food</td>
<td>25.36</td>
<td>28.09</td>
<td>22.97</td>
<td>0.72</td>
<td>19.86</td>
<td>21.21</td>
<td>20.70</td>
<td>0.68</td>
</tr>
<tr>
<td>Cereal</td>
<td>7.32</td>
<td>5.84</td>
<td>5.82</td>
<td>0.53</td>
<td>9.01</td>
<td>7.87</td>
<td>8.01</td>
<td>0.53</td>
</tr>
<tr>
<td>Coffee</td>
<td>11.14</td>
<td>9.23</td>
<td>7.64</td>
<td>0.69</td>
<td>11.67</td>
<td>10.93</td>
<td>8.98</td>
<td>0.77</td>
</tr>
<tr>
<td>Crackers</td>
<td>5.11</td>
<td>3.66</td>
<td>3.87</td>
<td>0.67</td>
<td>4.59</td>
<td>3.62</td>
<td>3.71</td>
<td>0.72</td>
</tr>
<tr>
<td>Eggs</td>
<td>4.49</td>
<td>3.20</td>
<td>3.20</td>
<td>0.50</td>
<td>3.55</td>
<td>2.46</td>
<td>2.17</td>
<td>0.63</td>
</tr>
<tr>
<td>Milk</td>
<td>7.81</td>
<td>6.32</td>
<td>6.51</td>
<td>0.27</td>
<td>6.63</td>
<td>5.35</td>
<td>5.16</td>
<td>0.37</td>
</tr>
<tr>
<td>Fresh meat</td>
<td>32.52</td>
<td>30.63</td>
<td>31.52</td>
<td>0.25</td>
<td>10.05</td>
<td>7.61</td>
<td>7.63</td>
<td>0.88</td>
</tr>
<tr>
<td>Detergent</td>
<td>10.23</td>
<td>8.49</td>
<td>9.63</td>
<td>0.69</td>
<td>7.95</td>
<td>7.15</td>
<td>7.78</td>
<td>0.75</td>
</tr>
<tr>
<td>Beer</td>
<td>26.32</td>
<td>23.24</td>
<td>21.61</td>
<td>0.83</td>
<td>24.05</td>
<td>23.73</td>
<td>19.24</td>
<td>0.91</td>
</tr>
<tr>
<td>Liquor</td>
<td>28.89</td>
<td>24.90</td>
<td>24.21</td>
<td>0.96</td>
<td>27.94</td>
<td>28.35</td>
<td>25.95</td>
<td>0.93</td>
</tr>
<tr>
<td>Fresh produce</td>
<td>23.82</td>
<td>22.39</td>
<td>24.50</td>
<td>0.17</td>
<td>9.19</td>
<td>9.00</td>
<td>8.91</td>
<td>0.41</td>
</tr>
<tr>
<td>Lawn and garden</td>
<td>29.65</td>
<td>51.24</td>
<td>20.49</td>
<td>0.89</td>
<td>11.79</td>
<td>12.59</td>
<td>10.48</td>
<td>0.94</td>
</tr>
<tr>
<td>Hair care products</td>
<td>13.06</td>
<td>16.65</td>
<td>9.51</td>
<td>0.79</td>
<td>8.04</td>
<td>7.76</td>
<td>6.93</td>
<td>0.82</td>
</tr>
<tr>
<td>Over the counter drugs</td>
<td>14.99</td>
<td>15.52</td>
<td>13.21</td>
<td>0.74</td>
<td>14.97</td>
<td>16.28</td>
<td>15.51</td>
<td>0.54</td>
</tr>
<tr>
<td>Oral hygiene</td>
<td>7.32</td>
<td>6.30</td>
<td>5.91</td>
<td>0.80</td>
<td>6.26</td>
<td>6.46</td>
<td>5.36</td>
<td>0.78</td>
</tr>
<tr>
<td>Shaving needs</td>
<td>11.47</td>
<td>10.91</td>
<td>10.36</td>
<td>0.94</td>
<td>8.55</td>
<td>9.53</td>
<td>8.40</td>
<td>0.93</td>
</tr>
<tr>
<td>Vitamins</td>
<td>24.83</td>
<td>33.05</td>
<td>17.15</td>
<td>0.91</td>
<td>19.26</td>
<td>19.48</td>
<td>17.91</td>
<td>0.79</td>
</tr>
<tr>
<td>Matched nondurables</td>
<td>239.20</td>
<td>181.94</td>
<td>219.81</td>
<td>0.07</td>
<td>151.47</td>
<td>117.47</td>
<td>143.00</td>
<td>0.10</td>
</tr>
<tr>
<td>Biweekly Observations</td>
<td>6,241</td>
<td>1,199,031</td>
<td>1,199,031</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Columns 1 and 5 show the mean of spending in the CEX Diary Survey and AC Nielsen, respectively, conditional on making a purchase, over a biweekly period in 2014. Columns 2 and 6 show the standard deviation of this spending across households. Columns 3 and 7 show the interquartile range (IQR) of this spending across households. Columns 4 and 8 show the zero share of spending on the specified category in the biweekly period in 2014. For the CEX Diary Survey, the sample of households is restricted to households reporting two diary weeks. By construction, the Diary Survey has no household with 0 spending in the biweekly period. For AC Nielsen, the sample of households includes only households with at least one shopping trip in each month of 2014. We aggregate daily spending to the biweekly period (weeks 1 and 2 of 2014 are one biweekly period, weeks 3 and 4 are a biweekly period, etc.) and treat the data as repeated cross-sections when calculating moments.
Average spending conditional on having a shopping trip during the biweekly period is higher in the CEX than Nielsen data. For example, the average total biweekly spending on comparable categories of nondurable goods is $239 in the CEX data compared with $151 in the Nielsen data. This difference reflects the fact that the Nielsen data report considerably lower levels of spending for categories with few UPCs such as “fresh meat,” “fresh produce,” “bread and baked goods,” and “lawn and garden.” The levels of spending are much closer for categories of goods populated by products with UPCs. For example, the average spending on “baby food” (a category where most goods have UPCs) in the Nielsen data is close to the average spending on “baby food” in the CEX data, while the average spending on “fresh meat” in the Nielsen data is smaller than the average spending on “fresh meat” in the CEX data. Note, however, that despite this limitation, the correlation between average spending or dispersion of spending in the CEX and Nielsen data is above 0.85 and can be further increased (up to 0.95) if a few outlier categories such as “fresh meat” are excluded. Measures of dispersion across the sources are close to each other too. We conclude that Nielsen data provide a useful complement to the CEX Diary Survey data for an analysis focusing on nondurable goods.

B. Evidence from the Volume/Size of Purchases

We can assess the possibility of increased stocking up by looking directly at volumes purchased; we should see not only increased spending per trip, but also increased physical volumes of goods purchased by households. Because Nielsen data report not only dollar spending for each UPC but also units purchased as well as volumes of units, we can check if this prediction is borne out by the data. In particular, the Nielsen dictionary of UPCs specifies count or weight for each UPC. Using 2004 as the benchmark year, we examine the distribution of purchased weights or counts for each module of goods. We identify a purchase as “large volume” if the purchased weight or count is greater than the ninetieth percentile of the size distribution of purchased weights or counts in the module. Then for each year we compute the share of purchases (by weight or count) for each module and household holding the “large volume” threshold constant across years. Using annual expenditure shares to aggregate across modules and sampling weights to aggregate across households, we construct an average share of “large volume” purchases for each year. We find that over time “large volume” purchases were increasingly prevalent. For products sold by weight, the share of large-volume purchases increased from 16.8 percent in 2004 to 22.3 percent in 2014. For products sold by count, the share rose from 19.3 percent to 21.8 percent over the same period. As before, these results are robust to conditioning on observable household characteristics or considering other thresholds for large-volume purchases (see online Appendix Figure A7). They

13 A module in the Nielsen data is a highly disaggregated category of goods. There are over 1,000 modules in the data. Example of modules are “FRUIT JUICE - APPLE,” “FRUIT JUICE - GRAPE,” “MEXICAN SHELLS,” “MEXICAN TORTILLAS,” “DAIRY-MILK-REFRIGERATED,” “DAIRY-BUTTERMILK-REFRIGERATED,” “DAIRY-CREAM-REFRIGERATED,” “EGGNOG - FRESH & CANNED.”

14 Results are similar when we consider alternative thresholds.
therefore provide direct evidence that households are engaging in larger-sized purchases of goods.

C. Evidence from the Intensive and Extensive Margins of Purchases

In addition to the evidence from inequality measures and days with positive shopping, we can assess more directly whether households are indeed stocking up more on their increasingly infrequent shopping trips. One indicator of shopping becoming increasingly concentrated over time would be that the amount of spending per shopping trip should be increasing as long as the total annual spending is stable. To assess this prediction, we decompose average annual expenditures by households into their average expenditure per shopping trip (the intensive margin) and their average number of shopping trips per year (the extensive margin). The results are presented in Figure 4. The figure shows six lines: average log annual spending per household, the average number of shopping trips per year, and the average log spending per shopping trip, and corresponding residualized series controlling for household observables, as in Section I. All series are normalized to be equal to one in 2004. While annual spending is approximately constant over 2004–2014, we see that the number of shopping trips declines by close to 20 percent while the average spending per trip increases by the same amount. Hence, households are doing fewer shopping trips, which is consistent with the CEX data for 1980–2015, but spending more on each trip. Furthermore, the results are unchanged if we control
for household observables, so these patterns are not driven by a changing composition of households but rather by changing behavior of households.

D. Evidence from Time Spent Shopping

Our argument suggests that households should increasingly buy goods in bulk and consequently spend less time shopping. While Nielsen data do not permit us to assess changes in shopping time for purchases of goods (e.g., we know the number of shopping trips but not their duration), we can use the American Time Use Survey (ATUS) to examine the evolution of households’ shopping time.

Since 2003, US Census Bureau on behalf of the Bureau of Labor Statistics (BLS) surveys a randomly chosen subset of households participating in the Current Population Survey (CPS) to report their time use for a given day. Each year, approximately 25,000 households are requested to recall their activities for a 24-hour period and provide detailed information on the type and duration of each activity. We use the American Time Use Survey (ATUS) to examine the evolution of households’ shopping time. Since 2003, US Census Bureau on behalf of the Bureau of Labor Statistics (BLS) surveys a randomly chosen subset of households participating in the Current Population Survey (CPS) to report their time use for a given day. Each year, approximately 25,000 households are requested to recall their activities for a 24-hour period and provide detailed information on the type and duration of each activity. We use the American Time Use Survey (ATUS) to examine the evolution of households’ shopping time.

For each activity start/end times are indicated, which allows us to observe how

---

15 There are precursors of the ATUS. An early time-use survey was implemented in 1965. Subsequent time-use surveys were done in 1975, in the mid-1980s and in the mid-1990s. Unfortunately, these earlier surveys differ in
many shopping trips were done by a respondent. Time spent for purchases of goods includes not only shopping time but also travel time, researching time, comparison time, etc. Because ATUS respondents are sampled from the CPS, we also have detailed demographic information (age, gender, marital status, educational attainment, employment status, income bracket, etc.).

Using this information, we compute average shopping time for each year and report the resulting series in Figure 5. In addition to the total time spent on purchases of goods, we show the intensive (average time per shopping trip) and extensive sample design, coverage, and level of detail. To ensure consistency of the series, we restrict our analysis to the surveys implemented by the BLS since 2003.

Figure 5. Shopping Time, American Time Use Survey

Notes: Panel A reports total shopping time (includes travel and other purchase related activities). Panel B reports the number of shopping trips per day conditional on having a shopping trip. Panel C reports the probability of having a shopping trip on a given day. Panel D reports the average duration of a shopping trip (including travel time and other purchase related activities; conditional on having a shopping trip). The black, solid line shows the raw average. The black, dashed line shows the average (regression) adjusted to demographic changes. See Section II for details.
(number of shopping trips per day) margins of shopping. Because the composition of the US population has been changing over time, we also present series adjusted for the changes using a specification similar to regression (1). The type of regression (1) depends on the nature of the dependent variable. When the dependent variable is the number of trips, we use a Poisson regression. For the average time spent on shopping for purchases of goods, we use a Tobit regression (because the distribution is censored at zero). For the average duration of shopping trips (which is conditional on having a trip), we use OLS. When the dependent variable is an indicator variable for having a shopping trip on a given day, we use a logit regression. In cases other than OLS, we take $\hat{\beta}$ as the mean partial effect.\footnote{In addition to demographic characteristics of households, these regressions include a set of dummies to capture within-week variation in shopping intensity and income brackets. We include dummies for week days because the day-of-week sampling of respondents has changed over time. We include controls for income because as income increases households may engage in more shopping (shopping is leisure) or less shopping (shopping is home production). We generally find that in the cross section, high incomes are associated with higher shopping time. Because buying smaller versus larger quantities of the same product should approximately take the same amount of time, one would expect the intensive margin not to vary as stocking up behavior changes although small variations could still occur if, e.g., consumers replace trips to multiple stores to purchase different products with trips to a single store where all products are sold.}

Figure 5 documents that shopping time (panel A) has been declining since 2003. Adjusting for the observed characteristics of respondents yields an even greater decrease. Panels B through D show that this reduction in shopping time is driven exclusively by the extensive margin rather than the intensive margin\footnote{The cumulative decline in the number of shopping trips is lower in the ATUS data than in the Nielsen data. This difference likely reflects the fact that the definition of a shopping trip is broader in the ATUS data than in the Nielsen data and covers goods with fewer opportunities to buy in bulk at low unit prices.}. Indeed, the average duration of a shopping trip (panel D) varies over time but does not exhibit any trend. In contrast, the probability of having a shopping trip (panel C) and the number of trips (panel B) decline over time.\footnote{Because buying smaller versus larger quantities of the same product should approximately take the same amount of time, one would expect the intensive margin not to vary as stocking up behavior changes although small variations could still occur if, e.g., consumers replace trips to multiple stores to purchase different products with trips to a single store where all products are sold.} These patterns are consistent with households doing fewer shopping trips but increasing the sizes of the products they buy during these trips. We find similar patterns for demographic subsamples (see online Appendix Figure A5). In addition, the decline in shopping trips with no rise in shopping time (panel D) suggests that households are engaging in larger purchases at approximately the same number of stores, not combining multiple store visits into single trips. As a result, it is unlikely that an increasing geographical concentration of stores into strip malls, shopping centers, etc., that lowers the fixed costs associated with a shopping trip can account for increased lumpiness of consumer purchases.

III. Quantifying the Contribution of Changing Shopping Patterns

Given this evidence on how shopping patterns have changed over time, we are interested in quantifying the potential contribution of this channel to explaining the differential trends that we observe across survey measures of expenditure inequality. We consider two ways of doing so. The first exploits the fact that, in the Nielsen data, we track high-frequency expenditures of households over extended periods of time and can therefore construct measures of expenditure dispersion that mimic the timing of the CEX Interview and Diary surveys, as well as even shorter and longer periods
of aggregation. This provides a direct test of how the frequency of aggregation can affect measured trends in expenditure inequality. The second method relies on the fact that we can also measure the dispersion in an individual’s expenditures over time, which will be directly related to their frequency of shopping. We show using a simple model that this “time-dispersion” in expenditures can contribute to the measured cross-sectional inequality in expenditures and that it can be used to quantify the contribution of changes in shopping patterns to trends in cross-sectional expenditure inequality.

A. Trends in Expenditure Inequality

As previously discussed, although we are limited to focusing on the period since 2004, when consumption inequality has diverged less dramatically across the CEX DS and IS, a key advantage of the Nielsen data is that we can examine directly how spending inequality varies with the level of time aggregation for the same set of individuals. Let \( X_{htpl} \) be spending of household \( h \) in period \( p \) (a week, bi-week, month, quarter of a given year, or a year itself) of calendar year \( t \) in location \( l \) (zip code, metropolitan area, or national level). Suppose the frequency of \( p \) is set to, e.g., a biweekly period. Then for each year \( t \), we calculate the cross-sectional coefficient of variation \( CV_tl = \sigma_{tl}/\bar{X}_{tl} \), where average spending for year \( t \) at frequency \( p \) over all households is \( \bar{X}_{tl} = (1/#h) \sum_h (1/#p) \sum_p X_{htpl} \) and the standard deviation of spending for the year \( t \) is \( \sigma^2_{tl} = (1/(#h \times p)) \sum_{hp} (X_{htpl} - \bar{X}_{tl})^2 \). The procedure for other frequencies is similar. In the Nielsen data, we treat weeks with no shopping activity over \( p \) as a true “zero”; in the DS and IS, we only keep households that have complete records over the length of \( p \) (two weeks of diaries for analysis at biweekly frequency, and four quarters of interviews for analysis at the annual frequency). Our measure of spending in the Nielsen data includes three major categories of goods: food-at-home, alcohol/tobacco, and small nondurables (e.g., paper towels, razors).

Panel A of Figure 6 plots the resulting measures of consumption inequality using the different levels of time aggregation for the 2004–2014 period. We observe two important patterns in the data. First, as we increase the level of time aggregation, the level of spending inequality declines. For example, the coefficient of variation at the biweekly frequency is approximately 0.8 but is less than 0.6 at the annual frequency. If household consumption were equal to household spending, we would not observe such differences. The fact that inequality of spending is decreasing in the level of aggregation is consistent with consumption being smoother than spending.

Second, the trends in expenditure inequality are different across frequencies. While spending inequality measured at high frequencies (e.g., biweekly) increases over time, it is much flatter when measured at low frequencies (quarterly and annual). Table 2 reports the average annual change in inequality by frequency and documents a sizeable, statistically significant decrease in the slope of the time trend until we reach the quarterly frequency of aggregation. Thus, simply changing the

\[19\text{ Because of the large sample sizes in the Nielsen data, bootstrapped standard errors on these series are close to 0 and are omitted.}\]
Figure 6. Inequality in Spending by Frequency of Time Aggregation, AC Nielsen Household Panel

Notes: The figures plot coefficients of variation (CV) of household expenditures when average expenditures are measured at different frequencies of time aggregation ranging from weekly to annually. The left panel shows the average annual cross-sectional coefficient of variation of expenditures across households, where expenditures are measured at different time frequencies. The right panel shows the average (across households) time-series coefficient of variation of each household’s expenditures over the course of the year, measuring expenditures at different frequencies. All calculations are for the Nielsen data. See Section III for details. Because of the large sample sizes in the Nielsen data, bootstrapped standard errors on these series are close to 0 and are omitted.

Table 2—Time Trends in Expenditure Inequality by Time Aggregation

<table>
<thead>
<tr>
<th>Dependent variable: Coefficient of variation</th>
<th>Frequency of aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Biweekly (1)</td>
</tr>
<tr>
<td><strong>Panel A. Nielsen data, 2004–2014</strong></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>0.0154</td>
</tr>
<tr>
<td></td>
<td>(0.0025)</td>
</tr>
<tr>
<td>Observations</td>
<td>11</td>
</tr>
<tr>
<td><strong>Panel B. CEX data (all nondurables), 1980–2015</strong></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>0.0048</td>
</tr>
<tr>
<td></td>
<td>(0.0009)</td>
</tr>
<tr>
<td>Observations</td>
<td>36</td>
</tr>
<tr>
<td><strong>Panel C. CEX data (nondurables as in the Nielsen data), 1980–2015</strong></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>0.0020</td>
</tr>
<tr>
<td></td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Observations</td>
<td>36</td>
</tr>
</tbody>
</table>

Notes: The table reports the estimated slope in the regression of the coefficient of variation for a given frequency of time aggregation on a trend. Time aggregation is indicated in the top row. Panel A uses data from AC Nielsen. Panel B uses CEX data covering all nondurable goods and services: the Diary Survey for column 1 and the Interview Survey of columns 3 and 4. For the Interview Survey of the CEX, the dependent variable in column 3 includes some expenditures that are measured at the monthly frequency. Panel C restrict the CEX data to cover only goods included in the Nielsen sample (food-at-home, alcohol/tobacco, and small nondurables). Newey-West standard errors are reported in parentheses.
time horizon over which one measures expenditures significantly alters the measured growth in expenditure inequality, and in precisely the direction that we would expect if households are reducing the frequency at which they purchase goods. This difference in time aggregation could potentially account for much of the difference in observed trends between the Interview and Diary survey measures of expenditure inequality.

We can use disaggregated data to further explore this insight. Specifically, for each module in the Nielsen data for year 2014, we compute the ratio of spending inequality at the biweekly frequency to spending inequality at the annual frequency. Then we relate this ratio to the frequency of shopping trips households have on average for goods in the corresponding modules.\(^{20}\) We find a strong negative relationship between the ratio and the frequency of shopping, which is consistent with the predictions of our theory. This figure also provides some insight into consumption versus expenditure inequality. Consider spending on “Men’s Toiletries,” which appears at the upper left of the figure. The largest component of this category is shaving products, which for most men are “consumed” on a daily basis, even though expenditure is infrequent (even in their smallest packages, razors and shaving cream will last for more than one use). This example illustrates that it is indeed

\(^{20}\)This frequency is taken from Baker and Kueng (2021). We are grateful to Lorenz Kueng for sharing the data.
the frequency of purchases that drives our results rather than change in consumption flows.

Given that the frequency of shopping trips has declined, we can, in principle, reconcile why the levels and trends are different for spending inequality measured at different frequencies. Dynamics of the ratio of CVs in the CEX data are consistent with this hypothesis. Specifically, we find (online Appendix Figure A9) that the ratio for the “food at home” category (many goods in this category are storable and can be purchased in bulk) rises more sharply than the ratio for “purchased meals” (effectively, non-storable goods). More generally, we find that inequality of spending on non-storable goods and frequently purchased services has similar trends in the IS and DS surveys (online Appendix Figure A8).

We can also directly explore the importance of the frequency of time aggregation for expenditures in the CEX data, albeit in a more limited way than in the Nielsen data. Column 1 of panel B in Table 2 reports the results for the growth in expenditure inequality in the Diary Survey. Within the Interview Survey, we can compare trend growth in expenditure inequality measuring expenditures at the quarterly (three-month period) frequency versus the annual frequency. Columns 3 and 4 of panel B in Table 2 report the results. As with the Nielsen data, growth is significantly larger when expenditures are measured at biweekly frequency, and we find no significant difference in the slopes at quarterly or annual frequency, suggesting that few purchases in these data are conducted at a less than quarterly frequency. We reach the same conclusions when we restrict the CEX data to include goods that are covered in our Nielsen sample (panel C). In short, the Nielsen data provide additional evidence that most of the difference in expenditure inequality trends observed between the Interview and Diary surveys can be accounted for by time aggregation of expenditures.

**B. The Dispersion over Time of a Household’s Expenditures**

A related approach to quantifying the contribution of time aggregation of expenditures to cross-sectional inequality is to consider the time-dispersion of households’ expenditures. When households make their purchases less frequently, we will observe rising dispersion in a household’s expenditures as we decrease the time horizon over which an individual’s expenditures are observed. Hence, we should observe similar patterns in the time dispersion of expenditures across frequencies of aggregation within individuals as we do in the cross-sectional data.

To see more precisely how the time-series dispersion of expenditures for individuals relates to the cross-sectional dispersion of expenditures, consider an environment similar in spirit to the celebrated Baumol-Tobin model. Specifically, each household $h$ consumes a target dollar amount of consumption $C_h$ over a total period of time of length $T$ (e.g., a year where $T = 26$ bi-weeks). Suppose a household makes equally sized purchases only on $N_h$ periods out of the $T$. In a period when the household makes a purchase (which happens $N_h/T$ of the time), that purchase is $X_{h,t} = C_h/N_h$, while in other periods (the remaining $1 - N_h/T$ fraction of the time) its purchases are $X_{h,t} = 0$.

We assume households can smooth their consumption over time so that, regardless of $N_h$, their per-period consumption is $\bar{C}_h \equiv C_h/T$. In this case, the cross-sectional
average level of consumption across all $T$ periods is $E_h[\bar{C}_h] \equiv \bar{C}$ and the corresponding dispersion of consumption, as measured by the cross-sectional coefficient of variation is $CV_h(\bar{C}_h) \equiv \sqrt{\text{var}(\bar{C}_h)}/\bar{C}$, which we take as given. Note that, if one could measure the dispersion in consumption for a given household over each of the $T$ sub-periods, it would be identical to the dispersion in consumption over a longer time period because by assumption households can perfectly smooth their consumption flow. If we could measure consumption flow directly, the time horizon used for measuring those flows would not matter for the resulting measures of the cross-sectional dispersion of consumption.

To see the link between the cross-sectional dispersion in expenditures and the time-variation in each household’s expenditures, it’s helpful to start with the latter. Suppose we measure expenditures for each of the $T$ subperiods for household $h$. The average expenditures across $T$ subperiods for household $h$ is $E_T(X_{htl}) = (C_{h}/N_{h}) \times (N_{h}/T) + 0 \times (1 - (N_{h}/T)) = C_{h}/T \equiv \bar{X}_{h}$ and the variance of these expenditures for household $h$ over the $T$ periods is $\text{var}_T(X_{htl}) = (N_{h}/T) \times ((C_{h}/N_{h}) - \bar{X}_{h})^2 + (1 - (N_{h}/T))(0 - \bar{X}_{h})^2 = \bar{X}_{h}^2(T/N_{h} - 1)$. Hence, the coefficient of variation over time for household $h$ when its expenditures are measured over subperiods is given by $CV_T(X_{htl}) \equiv \sqrt{\text{var}_T(X_{htl})}/\bar{X}_{h} = \sqrt{T/N_{h} - 1}$ so that a household’s time dispersion in expenditures is directly related to its frequency of shopping. As a household increasingly bunches its expenditures into fewer shopping trips ($N_{h}$ falls), the time-series dispersion in the household’s measured expenditures will rise.

To measure time-series variation in purchases in the Nielsen data, we follow our previous notation and let $X_{htpl}$ be spending of household $h$ in period $p$ (a bi-week, month, or quarter of a given year) of calendar year $t$ in location $l$. We calculate the average per period spending for household $h$ in year $t$ as $\bar{X}_{htl} = (1/#p)\sum_{p\in l}X_{htpl}$ and the variance of spending for household $h$ in year $t$ as $\sigma_{htl}^2 = (1/#p)\sum_{p\in l}(X_{htpl} - \bar{X}_{htl})^2$ and compute each household’s coefficient of variation for spending over the course of the year as $CV_{htl} = \sigma_{htl}/\bar{X}_{htl}$. Households who do more infrequent shopping trips have relatively higher standard deviation in their spending at higher frequencies than at lower frequencies. This greater time series dispersion in spending for one household when they do large purchases infrequently is therefore analogous to how cross-sectional dispersion in consumption is higher the more households engage in infrequent shopping. We calculate $CV$ at three frequencies: biweekly, monthly, and quarterly.

In panel B of Figure 6, we plot time series of the average (across households) coefficient of variation of each household’s expenditures over time (that is, $CV_{t} = (1/#h)\sum_{h}CV_{ht}$) using different time frequencies ranging from biweekly ($T = 26$ bi-weeks) to quarterly ($T = 4$ quarters); because of the large sample sizes in the Nielsen data, bootstrapped standard errors on these series are close to 0. As expected from infrequent purchases, the dispersion in households’ expenditures is higher on average at high frequencies of aggregation and as the frequency of aggregation declines, the measured dispersion in expenditures falls. For

\[^{21} \text{In the Nielsen sample design, households exiting the sample do not have observations in the last few days of the calendar year. To ensure this does not affect our results, we focus on the first 50 complete weeks of the year.} \]
example, the average time-series CV at biweekly frequency is about 0.7, while the average time-series CV at the monthly frequency is approximately 0.4. Importantly, we can see a rising trend in the time dispersion of households’ expenditures at the biweekly frequency, which is consistent with what we would expect when \( N \) is falling over time. The results are similar when we control for household characteristics.

To relate the cross-sectional and time-series measures of dispersion, note first that the coefficient of variation in expenditures summed across all \( T \) periods is equal to the cross-sectional dispersion in consumption: 

\[
CV_h(\bar{X}_h) = CV_h(\bar{C}_h).
\]

That is, by using a long period over which to aggregate expenditures, one can recover the underlying dispersion in consumption. Now suppose, instead, we can only measure expenditures each subperiod. Then one can show (see online Appendix C for derivations) that, under general conditions (e.g., goods may or may not depreciate), the cross-sectional coefficient of variation for expenditures across households at a given frequency is

\[
CV_h(X_{h,t}) \approx CV_h(\bar{X}_h) \sqrt{1 + \left( \frac{1}{CV_h(\bar{X}_h)^2} + 1 \right) \left\{ CV_T(X_{h,t}) \right\}^2},
\]

where \( CV_T(X_{h,t}) = E_T CV_h(X_{h,t}) \) is the average across households of the time dispersion of expenditures for each household at the given frequency. The first term \( CV_h(\bar{X}_h) \) captures the fact that higher underlying cross-sectional dispersion in consumption will lead to a higher dispersion of measured expenditures, even when the latter are measured at a higher frequency. For this term, we use \( CV_h(\bar{X}_h) \) with \( \bar{X}_h \) measured over a long horizon (e.g., a year) to measure \( CV_h(\bar{C}_h) \). The last term \( CV_T(X_{h,t}) \) reflects the fact that more dispersion in each individual’s expenditures over time will lead to a higher level of dispersion in the cross-section as well, as long as \( N_{h} < T \) for some \( h \), i.e., as long as households’ purchases are made less frequently on average than the length of the measurement period. As can be seen in Figure 6, increasing the duration over which expenditures are aggregated pushes the time dispersion of expenditures toward zero, but the cross-sectional dispersion is converging to a positive value, which according to the model is equal to the underlying dispersion of consumption.

Using average values of \( CV_T(X_{h,t}) \) at the biweekly frequency and \( CV_h(\bar{X}_h) \) at the annual frequency in the Nielsen data yields

\[
\frac{\partial CV(\bar{X}_{h,t})}{\partial CV_T(X_{h,t})} \approx 1.22
\]

Given that the time-series dispersion in expenditures has risen by about the same amount as the cross-sectional dispersion between 2004 and 2014 at the biweekly frequency implies that the decreased frequency of expenditures can account for all of the rise in inequality of expenditures across households at the biweekly frequency of aggregation in the Nielsen data.

\[\text{22 With those same parameter values, we can also verify that } \frac{\partial CV(X_{h,t})}{\partial CV_T(X_{h,t})} \approx 1 \text{ so changes in underlying consumption inequality translate one-for-one into changes in expenditure inequality measured at biweekly frequency of expenditure aggregation.}\]
IV. The Rise of Club Stores and Expenditure Inequality

Previous sections document that US households spend less time shopping and make fewer shopping trips, so that inequality of expenditure measured at high frequency can rise over time while inequality of consumption can remain stable. Obviously, there are many possible sources underlying this changing behavior of US households, but one such mechanism is likely the rise of club (warehouse) stores (e.g., Costco, Sam’s Club, BJ’s) which, by design, sell larger quantities of goods to households at lower unit prices and encourage households to buy goods in bulk. As a result, it has become easier for households to stock up in ways that were not feasible in the past.

To assess the role of club stores, we collected data on BJ’s, Sam’s Club, and Costco store openings by contacting these companies for lists of stores and conducting internet searches for store opening dates.\(^{23}\) Indeed, club stores have expanded dramatically throughout the country since the 1980s (see panel A of Figure 8), which is consistent with the observed trend in expenditure inequality. To measure intensity of shopping in club stores for a given year, we use the fraction of a household’s expenditures that was spent at club stores over the course of that year. Specifically, we calculate the share as

\[
\text{share}_{htl}^{(\text{club})} = \frac{\sum_{p \in t} X_{htpl}^{(\text{club})}}{\sum_{p \in t} X_{htpl}},
\]

where \(X_{htpl}^{(\text{club})}\) is spending at club stores. In the Nielsen data, the share of household spending at club stores in spending on goods in our sample (food-at-home, alcohol/tobacco, and small nondurables) increased from 7.9 percent (\(\approx$320 per year\)) in 2004 to 9.8 percent (\(\approx$390 per year\)) in 2014 (see panel B of Figure 8).\(^{24}\) Panel C of Figure 8 shows the degree of market penetration of these stores by plotting the distribution of households in the Nielsen data and their distance from the nearest warehouse/club store (in 2004 and 2014). While there is considerable variation in the ease with which households can access one of these retailers, approximately 40 percent of households live less than 5 miles away from one of these stores. At the same time, 30 percent of households have to drive more than 10 miles to reach the nearest store and almost 20 percent have to drive 25 or more miles. Panel D shows the geographical distribution of club stores in 2005.

To assess whether club stores can explain some of the rising concentration in household shopping trips, we characterize the link between how much variation there is in a household’s expenditure over time and their reliance on club stores. We do so by regressing households’ time-series coefficients of variation on households’ club share expenditures, using coefficients of variation measured at different time frequencies. In other words, we estimate the following specification:

\[
CV_{htl} = \beta \times \text{share}_{htl}^{(\text{club})} + X_{htl} \gamma + \lambda_t + \psi_h + \text{error},
\]

\(^{23}\) If the Sam’s club was first a PACE Membership Warehouse club, we date the Sam’s club opening date as the opening of the PACE club.

\(^{24}\) We exclude durables in Nielsen from this analysis since club stores also sell durables—although not in bulk—which would drive up the club share; however, our results are also robust to including durables.
where $\lambda_t$ and $\psi_h$ are the year and household fixed effects, and $\mathbf{X}$ is a vector of controls (the number of children, female head of households, employment status, income brackets, race, employment status of household head, educational attainment of household head, age, and age squared for household head). We use information in $\mathbf{X}$ to control for changes in household characteristics over time (e.g., greater participation of women in labor force, rising incomes, aging of population). To make inference conservative, we cluster standard errors at the zip-3 level (i.e., first three digits of zip code).

Our theory predicts a positive relationship between time-series $C\text{V}_{htl}$ and $\text{share}_{htl}^{(\text{club})}$: as a household buys a greater share of their budget at club stores, their purchases should be lumpier. However, causation could run in the opposite direction. For example, if some households choose to have significant time variation in their expenditures (for example, because they like to host a party every month), they might also be more likely to go to club stores to stock up for these events. To rule out this alternative direction of causality, we pursue an instrumental variable approach in which our instrument is proximity to a club store (as measured...
This exploits time-series variation, e.g., stores open and reduce the distance to the nearest club stores faced by some households. Our identifying assumption is that the timing of the store opening is exogenous from the perspective of the household. It is very likely club stores are choosing their location to maximize the size of their customer base. Our household fixed effects and time-varying demographic controls should address any targeting on observable characteristics. In online Appendix Table A1, we document which household characteristics are strong predictors of shopping at club stores. Still, if a club store enters in years when local household demand changes and households would have otherwise travelled further to access club stores, or households anticipate these openings by changing spending patterns in dramatic ways, then this would be a threat to our identification. To strengthen the quality of our instrumental variables, we exclude households who moved from one location to another. As a result of this restriction, time variation in distance to a club store is determined exclusively by entry/exit of stores.

To construct a measure of distance from club stores, we created a database of geographical locations and openings/closures of club stores for the three largest chains: Sam’s Club, Costco, and BJ’s. For example, we know that the Costco store in Richmond, California was opened on October 16, 1986, at 4801 Central Avenue. A household’s distance from the nearest club store is calculated between the centroid of the zip code where a given household lives and the centroid of the zip code of the club store.

The results (Table 3) suggest that shopping at club stores is indeed significantly correlated with more stocking up. First, looking at high frequencies, such as weekly, there is a positive statistically significant coefficient on the share of expenditures going to club stores, so households that spend relatively more at these stores display more volatility in their expenditures across weeks in a year. However, when we increase the time span over which expenditures are measured, this coefficient shrinks rapidly. At the quarterly frequency, shopping at club stores leads to much less time variation in quarterly spending, which is as expected since it becomes progressively more difficult to stock up for longer periods. Shopping at club stores also explains a diminishing fraction of the variance in households’ coefficients of variation at longer durations for time aggregation of expenditures.

Table 3 also shows that the distance to a club store is a strong instrument for the share of spending at club stores in total spending. Households located further from club stores display significantly smaller shares of expenditures at these stores. The first stage F-statistic is above 30. Overall, the OLS and IV estimates paint the same picture. This finding supports the notion that the rising access to club stores has induced households to increasingly stock up on goods and reduce the frequency of

---

25 Club stores are also known for having lower unit prices. The effects on consumption and expenditure might further diverge if households who access club stores do so at lower unit prices. As a robustness check to address this concern, we replaced unit prices at club stores with average unit prices outside of club stores for the same brand (e.g., “Bounty Paper Towels”). We recalculate the CV and rerun the regressions reported in Table 3. We find qualitatively similar results (online Appendix Table A2), suggesting that any price differences at club stores are not driving our results.

26 Consistent with these results, we also find (online Appendix Figure A10) that households have longer intervals between shopping trips to club (warehouse) stores than intervals between shopping trips for other stores.
their shopping trips. In turn, this change in shopping behavior has generated spending patterns that appear more unequal in the cross-section when measured at high frequencies, even if their underlying consumption flows have not changed.

The results are even stronger if we restrict our attention to households with two or more members (panel B of Table 3), whereas the effects are quite small for single-member households. This is consistent with the idea that club stores are more advantageous for larger households, whereas single-member households may find it less useful to purchase very large quantities of each type of good. In online Appendix Table A1, we document a number of other household characteristics that are associated with higher club store spending, such as education and income. However, we focus on results across all households since we map our estimates into cross-sectional dispersion measures that include all households.

Quantitatively, these estimates are economically significant. From 1980 (a year with a negligible share of warehouse stores in total sales) to 2014, the average expenditure share of club stores has risen by approximately 10 percentage points. Given the IV estimates in Table 3 (panel A), this implies that club stores can single-handedly account for approximately 30 percent of the trend rise in the cross-sectional dispersion of expenditures measured at the biweekly frequency relative to the quarterly

<table>
<thead>
<tr>
<th>Dependent variable: Coefficient of variation</th>
<th>Frequency of aggregation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weekly</td>
</tr>
<tr>
<td>OLS (1)</td>
<td>OLS (2)</td>
</tr>
<tr>
<td>Club share</td>
<td>0.238</td>
</tr>
<tr>
<td>(0.012)</td>
<td>(0.222)</td>
</tr>
<tr>
<td>Observations</td>
<td>393,822</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.768</td>
</tr>
<tr>
<td>First stage $F$-statistic</td>
<td>38.14</td>
</tr>
</tbody>
</table>

Panel B. Families

| Club share                                 | 0.224  | 0.569   | 0.136   | 0.348   | 0.057   | 0.052  | 0.011   | 0.004  |
| (0.014)                                    | (0.215) | (0.011) | (0.163) | (0.008) | (0.128) | (0.006) | (0.096) |
| Observations                               | 290,742 | 290,742 | 290,742 | 290,742 | 290,742 | 290,742 | 290,742 |
| $R^2$                                      | 0.778  | 0.178   | 0.732   | 0.107   | 0.637   | 0.041  | 0.503   | 0.011  |
| First stage $F$-statistic                  | 37.45  | 37.45   | 37.45   | 37.45   | 37.45   | 37.45   | 37.45   |

Notes: The dependent variable is the coefficient of variation (CV) calculated as follows. For each household, we calculate (i) standard deviation of spending at a given frequency (weekly, biweekly, monthly, quarterly) for a given year and (ii) average spending per period (total annual spending divided by the number of periods with shopping trips). The coefficient of variation (CV) is (i) divided by (ii) so that CV is time-series volatility of spending for a given household in a given year. Club share is the share of annual spending at club stores (Sam’s Club, Costco, BJ’s, etc.) in total annual spending at all stores. Spending includes only food, alcohol/tobacco, and small nondurables (paper towels, toothpaste, etc.). The sample of households includes only households with at least one shopping trip in each month of a given year. For each household, the instrumental variable is the distance to the closest club store (Sam’s Club, Costco, BJ’s). This distance is calculated between the centroid of the zip code where a given household lives and the centroid of the zip code where the nearest club store is located. Regressions include but do not report coefficients on the following controls: year and household fixed effects, age and age squared for the household head, a set of dummy variables for household income brackets, number of children, employment status, race, educational attainment, gender of household head. Standard errors are clustered at the zip-3 level (i.e., first three digits of zip code). Panel A is for all households, while panel B restricts to households with a household size of at least two persons.
or annual frequency, since $0.1 \times 0.22$ implies a 0.022 contribution to the level of dispersion coming from club stores or equivalently 0.0007 per year, relative to an average rise of 0.002 (see column 1, panel C, Table 2) over the same time period in the cross-sectional coefficient of variation measured biweekly in the CEX for the same set of goods that we use in the Nielsen data.\footnote{We can map one-to-one from the effect on time dispersion to the effect on cross-sectional dispersion as shown in Section IIIB.} Note that this quantification of the club store contribution may underestimate the influence of club stores as these stores can influence the behavior of other stores, for example, by inducing convenience and department stores to start selling multipacks or large packages of goods to keep up with club stores.

While the increasing prevalence of club retailers appears to have contributed significantly to changing consumer shopping patterns, there are a number of other complementary explanations that could also help account for these trends. For example, anything raising the fixed cost of shopping trips, be it financial (e.g., rising gas prices, rising opportunity cost of time, increasingly moving away from city centers and stores) or in terms of the amount of time (e.g., through rising traffic), would contribute to the declining frequency of shopping. The decline in the real price of gasoline since the early 1980s suggests that gasoline prices are unlikely to have been an important contributor to this changing behavior. However, rising levels of traffic and the progressive “suburbanization” of US cities are more difficult to rule out given the data currently available.

Another force that could lead to a falling frequency of shopping is the decline in the cost of storage. For example, increased ownership of refrigerators/freezers has allowed for more storage of food products, but this is unlikely to be an important contributor since we observe increased stocking up across a wide range of goods, not just food products. The growing size of US houses, on the other hand, could induce more stocking up on the part on households.\footnote{Median square footage of US single-family homes has increased 12.5 percent since 1985, from 1,600 square feet in 1985 to 1,800 square feet by 2013 (authors calculation; data from the American Housing Survey).} In the absence of detailed information on the changing sizes of homes across regions, it is difficult to quantify this channel precisely with our data although this would be a promising area for future research.

A final force worth considering is financial innovation. Much of the work focusing on consumption inequality has explained the flat profile found in the Interview Survey through the financial innovation channel. According to the leading hypothesis, expanded access to credit has allowed households to better smooth transitory economic shocks, thereby pushing down consumption inequality, even though the prevalence of transitory shocks, reflected in rising income inequality, has been increasing. This same financial innovation expanding credit may also have allowed households to better take advantage of bulk discounts like those available at club/warehouse stores. Interestingly, the expansion of credit may have acted to raise spending inequality through our mechanism when looking at high-frequency shopping patterns, while reducing this inequality in lower frequency data by allowing households to mitigate transitory income shocks. In future work, we intend to use data as in Gelman et al. (2017) to examine the relationship between the availability of credit and bulk shopping.
V. Conclusion

There has been growing interest in the apparent difference in trend between expenditure and income inequality documented by Slesnick (2001) and Krueger and Perri (2005). Since then, much of the literature has focused on the difficulties associated with measuring expenditure inequality (specifically, underreporting of expenditures) and concluded that it has, in fact, increased in line with income inequality. We document another measurement issue with consumption measures, namely the infrequent timing of many expenditures, which suggests that consumption inequality has likely increased by less than standard measures imply.

Specifically, since households engage in infrequent purchases of many goods, when expenditures are measured at a high frequency many households will appear not to purchase these goods, leading to the appearance of high inequality in consumption, even though their consumption may in fact mirror that of households who are observed to purchase the good. We document that households are engaging in fewer shopping trips than in the past and buying larger volumes and quantities when they do make purchases. These trends will, when combined with high-frequency measures of expenditures, lead to the appearance of rising expenditure inequality even when none is present. We show that these patterns can account for much of the rise in expenditure inequality in the Diary Survey of the CEX and that a lower frequency of aggregation of expenditures points toward little change in consumption inequality. A major force behind this changing consumption behavior appears to be the rise of club/warehouse stores which facilitate and encourage larger sized purchases. As the market for club/warehouse stores becomes more saturated and as bulk goods become more prevalent even in non-club/warehouse stores, one may expect the patterns documented here to have less of an impact on measured spending inequality in the future.

Relatedly, the growing prevalence of online retailing and home deliveries is reducing fixed costs of shopping associated with low unit prices and should therefore be pushing toward a higher frequency of shopping. As online retailing continues to grow in both size and scope, this implies we may observe a reversal of some of the patterns documented here. It therefore seems promising for future work to consider how these different forces will balance out so that we can better understand how to properly measure underlying trends in consumption inequality.

Our analysis identifies an apparent trade-off in how spending data are currently collected. One option is to rely on interviews to gather information about a wide variety of spending categories over relatively long horizons (e.g., the Interview Survey in the CEX). However, this richness comes at a cost of measurement/recall/non-reporting errors, which can attenuate inequality in consumption and possibly cloud inference about trends in inequality. Indeed, statistical agencies have had an increasingly hard time recruiting households to participate in interview-based surveys, thus motivating the use of alternative sources. The other option is to measure detailed spending precisely over short horizons (e.g., the Diary Survey in the CEX, spending for a typical week in the Panel Study of Income Dynamics) to ensure that respondents are willing to participate in data collection. While the quality of these data is arguably better, the short duration of measurement makes the link between
spending and consumption less straightforward for goods and services purchased infrequently. Clearly, an ideal approach for measuring consumption inequality should combine the strengths of both options and minimize the intrusiveness of reporting. Perhaps, data collection similar to Gelman et al. (2017) provides a key component for future surveys of consumer expenditures.

When interpreting our results, one should bear in mind an important caveat. Our analysis focuses on groceries and small nondurables for which we have precise measurements of spending at various frequencies and which have been routinely used in previous analyses of consumption/spending inequality. Although we find that consumption inequality for these categories changed little since the 1980s, it remains to be seen whether this result generalizes to broader measures of consumption. For example, rising income inequality can translate into greater inequality for consumption of luxury-like goods (e.g., spas, travel, jet fuel, high-end durables, and housing). Given our data constraints, we will not be able to detect such a trend, but future work may have better data or use tools such as those in Aguiar and Bils (2015) to take advantage of accurate spending data collected at high frequencies to make inferences about the evolution of broader measures of consumption inequality.

This caveat also raises an important question about what measure of consumption inequality should be used for policymaking. Specifically, one may entertain the possibility that a reasonable policy objective could be to minimize inequality in the consumption of necessities so that basic needs are satisfied for a wide spectrum of population. However, it is also conceivable that inequality in conspicuous consumption may be particularly damaging for the cohesion of a society, and therefore policymakers should target inequality for total consumption. We hope that future theoretical work will provide more guidance on what measure empirical research should concentrate on.

REFERENCES


29 Interestingly, whether this additional burden translates into better measures of consumption is not well established. Bee et al. (2012) find evidence that spending in the Diary Survey compares less favorably to NIPA aggregates than does spending captured in the Interview Survey for certain categories of goods, suggesting other issues with the Diary Survey.


